Lassen sich ordinale Daten mit linearen Strukturgleichungsmodellen analysieren?

Kühnel, Steffen M.

Lassen sich ordinale Daten mit linearen Strukturgleichungsmodellen analysieren?

von Steffen Kühnel [1]

Zusammenfassung

Lineare Strukturgleichungsmodelle setzen voraus, daß die abhängigen Variablen metrisches Meßniveau haben. In der Praxis werden sie aber auch bei ordinalen Variablen eingesetzt. Lassen sich ordinale Daten als ungenaue Erfassung metrischer Größen auffassen, kann untersucht werden, welche Auswirkungen das Ignorieren des ordinalen Meßniveaus hat und ob die Berücksichtigung des Meßniveaus durch vorgelagerte Schwellenwertmodelle die Modellschätzungen verbessert. Anhand eines kleinen Monte-Carlo-Experiments wird gezeigt, daß trotz des Ignorierens des ordinalen Meßniveaus korrekte von fehlspezifizierten linearen Modellen unterschieden werden können. Wird die Ordinalität der Daten explizit berücksichtigt, verbessert sich die Genauigkeit der Parameterschätzungen. Diese Ergebnisse gelten auch dann, wenn die Schwellenwertparameter nicht konstant sind, sondern über die Untersuchungseinheiten zufällig variieren.

Abstract

The use of linear models presumes that dependent variables are metric. What will be the consequences, if the dependent variables are ordinal in the sense of rough categorical - measures of unobserved metric variables? A small Monte-Carlo study shows that it may be possible to discriminate between correct and mispecified linear models even if the ordinality of the data was ignored. Nevertheless, specifying the ordinal measurement process by threshold models may result in better parameter estimations. An interesting new conclusion of the Monte-Carlo study is that these results seems to hold even if the thresholds are not fixed but random.

1 *Dr. Steffen-M. Kühnel* ist Hochschulassistent am Institut für Angewandte Sozialforschung der Universität zu Köln, Greinstr. 2, D-50939 Köln. Der Beitrag basiert auf einem Vortrag anläßlich der Frühjahrstagung 1993 der Methodensektion der DGS in Mannheim.

Ein alter Einwand gegen den Einsatz von linearen Strukturgleichungsmodellen in der Sozialforschung bestellt darin, daß diese Modelle bei abhängigen Variablen metrisches Skalenniveau verlangen, in den Sozialwissenschaften in der Regel aber höchstens ordinale Messungen vorlägen. In der Vergangenheit haben die Proponenten linearer Modelle diesen Einwand meist mit der Bemerkung zurückgewiesen, daß das Ignorieren unterschiedlicher Abstände zwischen den Ausprägungen ordinaler Messungen relativ unproblematisch sei, wenn die ordinale Variable vier oder mehr Ausprägungen habe (vgl. ***Bentier und Chou*** 1987, S. 88). In der Praxis wurden lineare Strukturgleichungsmodelle aber auch bei dichotomen oder trichotomen abhängigen Variablen angewendet.

Die jüngste Generation von Programmen zur Analyse linearer Strukturgleichungsmodelle ermöglicht nun, durch vorgelagerte Schwellenwertmodelle die Ordinalität bzw. Nicht-Kontinuität von Variablen in linearen Modellen explizit zu berücksichtigen.[2] Wenn daraufhin von den Programmentwicklern die Beachtung des Meßniveaus gefordert wird[3], stellt sich allerdings schon die Frage, ob nicht in der Vergangenheit anstelle inhaltlicher Gründe eher die Möglichkeiten des eigenen Analyseansatzes die jeweilige Position bestimmten.

Ein Schwachpunkt mancher Diskussionen um die Beachtung des 'korrekten' Skalenniveaus ist meiner Ansicht nach, daß übersehen wird, daß gar keine Einigkeit darüber besteht, was ordinales Skalenniveau bei sozialwissenschaftlichen Daten bedeutet. Ich möchte diesen Beitrag daher zunächst mit einigen methodologischen Bemerkungen zum ordinalen Meßniveau beginnen. Dabei wird deutlich werden, daß im Kontext linearer Modelle 'ordinales Skalenniveau' eine spezifische Bedeutung hat. Vor dem Hintergrund dieser speziellen Sichtweise von Ordinalität werde ich dann auf Konsequenzen der Anwendung linearer Modelle bei ordinalen Daten eingehen.

2 Im LISREL-Programm ist die explizite Berücksichtigung der Ordinalität der Daten zwar bereits seit ungefähr 10 Jahren durch die Berechnung polychorischer Korrelationen möglich (ab LISREL 6), aber erst in *Muthéns* (1988) Programm LISCOMP wurde eine WLS-Schätzmethode implementiert, die auch inferenzstatistische Aussagen zuließ. Inzwischen sind diese Schätzverfahren auch in den beiden bekanntesten Programmen zur Analyse linearer Strukturgleichungsmodelle PRELIS 2/LISREL 8 *(Jöreskog und Sörbom* 1993a, b) und EQS 4 *(Bentier und Wu,* 1993) verfügbar.

3 So schlagen etwa *Jöreskog und Sörbom* (1993c, S. 45) vor, bei ordinalen Variablen statt Produkt-Moment-Korrelationen polychorische Korrelationen mit der WLS-Schätzmethode zu analysieren.

Ordinales Meßniveau in sozialwissenschaftlichen Daten

Meßwerte werden üblicherweise[4] als ordinal bezeichnet, wenn es möglich ist, die Meßobjekte hinsichtlich der gemessenen Eigenschaft in eine Rangordnung zu bringen. Für zwei beliebig herausgegriffene Meßobjekte kann dann festgestellt werden, ob bei einem der beiden Objekte die gemessene Eigenschaft weniger stark ausgeprägt ist als bei dem anderen Objekt, oder aber, ob sich die beiden Objekte in dieser Beziehung nicht unterscheiden. Falls sich nun zwei Objekte unterscheiden, liegen jedoch keine Informationen über das Ausmaß der Verschiedenheit vor, ist also unbekannt, ob der Unterschied zwischen den beiden Objekten eher klein oder eher groß ist.

Tabelle 1: Beispiel zur Konsequenz einer rangerhaltenden Datentransformation für den Vergleich von Mittelwerten

	ursprüngliche Meßwerte			transformierte Meßwerte		
	männlich		weiblich	männlich		weiblich
	1.0		1.0	1.0		1.0
	2.0		1.0	4.0		1.0
	3.0		2.0	9.0		4.0
	4.0		5.0	16.0		25.0
	5.0		5.0	25.0		25.0
$\bar{x}$:	3.0	>	2.8	11.0	<	11.2

Die Kritik an der Verwendung linearer Modelle bei ordinalen abhängigen Variablen beruht nun darauf, daß die Berechnung von Mittelwerten in linearen Modellen den Abstand von Meßwerten berücksichtigt. Wenn aber auf empirisch gar nicht vorliegende Informationen rekurriert wird, haben die Ergebnisse solcher Analysen keinerlei Aussagekraft. Dies läßt sich durch ein einfaches Beispiel demonstrieren, bei dem die ursprünglichen Meßwerte durch auf Ordinalskalenniveau zulässige Transformationen verändert werden. Obwohl die in den Daten enthaltene empirische Information bewahrt wird, sich die Rangfolge der Meßwerte also nicht ändert, kann der Vergleich zweier Mittelwerte vor und nach der Transformation zu entgegengesetzten Resultaten führen (vgl. Tabelle 1).[5]

4 Vgl. die entsprechenden Definitionen in Methodiklehrbüchern, z.B. in *Schnell u.a.* (1993, S. 149).

5 Dieses Beispiel orientiert sich an Vorbildern in *Kriz* (1981, S. 124f.) und *Kromrey* (1986, S. 122).

Anwender linearer Modelle stellen zwar nicht in Frage, daß bei ordinalen Daten die Abstände zwischen den Meßwerten unbekannt sind. Dies bedeute aber nicht, daß es tatsächlich vollkommen beliebig sei, welche Zahlenwerte die Meßwerte aufweisen, solange nur die empirisch gefundene Rangordnung der Meßobjekte gewahrt bleibe. In Unkenntnis der 'wahren' Abstände zwischen Meßwerten müsse man vielmehr diese entweder schätzen oder aber näherungsweise von etwa gleichen Abständen zwischen den Rangplätzen ausgehen. Ein grundsätzliches Verneinen der Anwendbarkeit linearer Modelle bei ordinalen Daten hieße dagegen, unnötig auf inhaltlich nutzbare Informationen zu verzichten. Letztlich liegt hier ein anderes Verständnis von Ordinalität vor: Während die Kritiker linearer Modelle davon ausgehen, daß bei ordinalem Meßniveau definitionsgemäß maximal Ranginformationen existieren, interpretieren die Befürworter dieser Modelle ordinale Messungen als ungenaue Erfassungen metrischer Größen.

Es macht m.E. wenig Sinn, die eine oder die andere Auffassung von Ordinalität als die einzig richtige zu bezeichnen. Stattdessen ist in jedem Einzelfall zu fragen, ob eine inhaltlich interessierende Größe aufgrund theoretischer Überlegungen als metrisch konzeptualisiert werden kann. Bei einer positiven Antwort kann dann versucht werden, Meß- und Analysemodelle anzuwenden, die die theoretisch vorhandenen Informationen möglichst weitgehend nutzen.

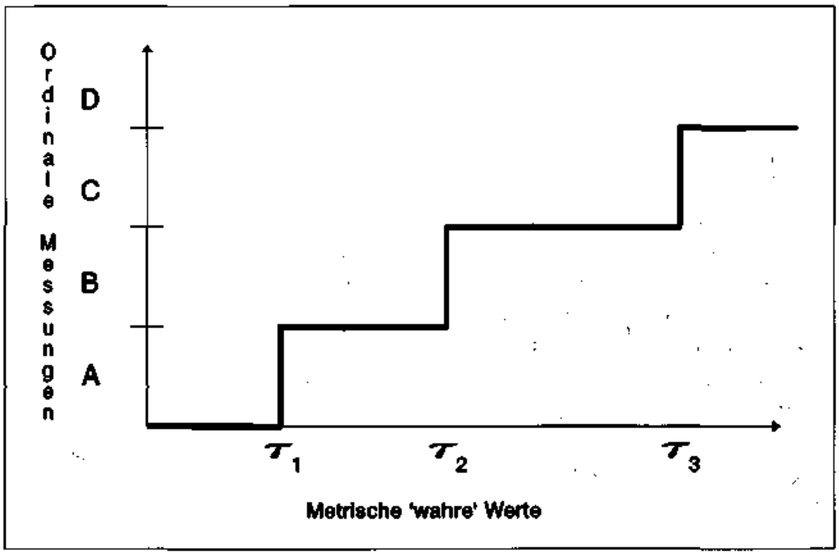
Bei den in der Umfrageforschung üblichen Einstellungs- und Meinungsfragen dürfte es im allgemeinen nicht schwer fallen, die Position eines Befragten als auf einer metrischen Skala liegend vorzustellen und die Erfassung über vorgegebene Antwortkategorien als eine ungenaue Messung einer eigentlich metrischen Größe aufzufassen. So läßt sich z.B. die Sympathiebeurteilung eines Politikers theoretisch auf einer unbeobachtbaren metrischen Sympathiedimension verorten. Wird nun ein Proband gebeten, einen Politiker auf einer vorgegebenen kategorialen Antwortskala zu beurteilen, dürfte die gewählte Antwortkategorie von der Positionierung auf der latenten metrischen Skala abhängen.

In einer solchen Situation ist es dann nicht von vornherein sinnlos, nach den Konsequenzen zu fragen, die sich ergeben, wenn man die Ungenauigkeit der Messungen ignoriert und die Abstände zwischen den Antwortkategorien als gleich betrachtet. Darüber hinaus läßt sich hier untersuchen, ob es den Aufwand lohnt, die Ungenauigkeit der Messungen explizit zu berücksichtigen. Die Beantwortung dieser beiden Fragen setzt ein Meßmodell voraus, daß die unterstellte Umfragesituation beschreibt, ordinale Antwortkategorien also als ungenaue Messungen metrischer Größen auffaßt. Im Zusammenhang mit linearen Strukturgleichungsmodellen werden die schon erwähnten Schwellenwertmodelle eingesetzt, deren Logik ich im nächsten Abschnitt vorstellen möchte.

Schwellenwertmodelle: Die Modellierung ordinaler Daten als grobe Messungen metrischer Größen

Die Interpretation ordinaler Messungen als ungenaue Erfassung einer eigentlich metrischen Größe läßt sich auf verschiedene Weise modellieren. Man kann etwa annehmen, daß die Kategorien der ordinalen Messungen diskreten Ausprägungen der metrischen Variable entsprechen. Unbekannt ist dann allein der 'wahre' Abstand zwischen den Meßwerten. Es ist aber auch vorstellbar, daß die metrische Variable kontinuierlich ist. Bei dieser Sichtweise entsprechen Teilbereiche des Wertebereichs der metrischen Variable den Kategorien der ordinalen Messungen.

Abbildung 1: Graphische Darstellung eines Schwellenwertmodells



In Abbildung 1 ist eine solche treppenförmige Beziehung zwischen den ordinalen Messungen und den metrischen 'wahren' Werten graphisch wiedergegeben. Ist der 'wahre' Wert kleiner als τ_1, entspricht dies bei der ordinalen Messung der ersten Kategorie (A). Metrische Werte zwischen τ_1 und τ_2 fallen in die zweite Kategorie (B), Werte zwischen τ_2 und τ_3 in die dritte Kategorie (C). Sind die Werte größer als τ_3, ergibt sich bei der ordinalen Messung die letzte Kategorie (D). Wie man sieht, erfolgt der Wechsel von einer Kategorie zur nächsten immer dann, wenn ein spezifischer Wert überschritten wird. Da diese Werte wie Schwellen wirken, die zur Erreichung einer Kategorie überschritten werden müssen, werden die Werte τ_i auch als Schwellenwerte (thresholds) und das Meßmodell selbst als Schwellenwertmodell bezeichnet.

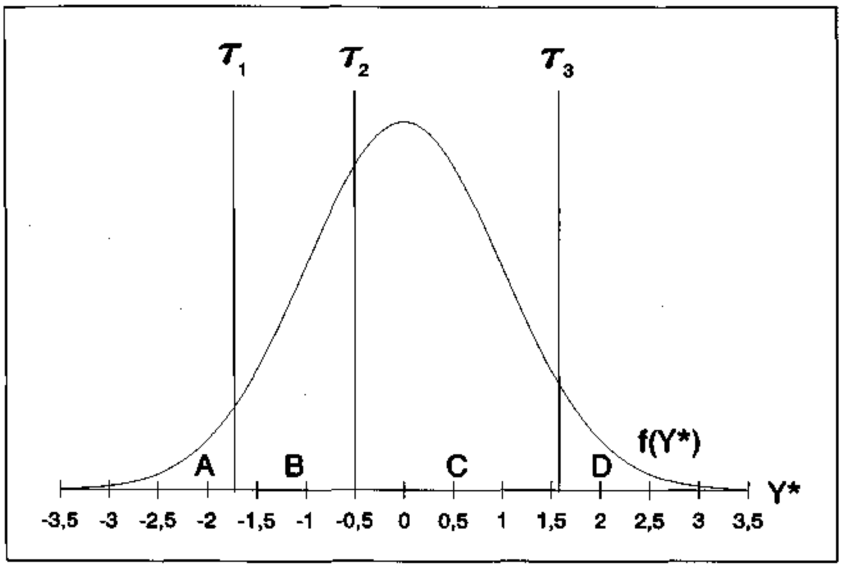
Schwellenwertmodelle ermöglichen die Untersuchung der Frage, ob und inwieweit die Mißachtung des ordinalen Meßniveaus Analyseergebnisse verfälschen kann. Mißachtung des Meßniveaus heißt hier, daß man die ordinalen Messungen als Realisierungen metrischer Variablen auffaßt und den ordinalen Kategorien aufeinanderfolgende (ganze) Zahlen zuordnet. Berechnet man nun für diese Zahlen erste und zweite Momente (Mittelwerte, Varianzen und Kovarianzen bzw. Korrelationen) oder nutzt sie zur Schätzung der Parameter linearer Modelle, so kann man die Abweichungen von den Werten der Momente oder Modellparameter betrachten, die man erhielte, wenn für die Berechnungen anstelle der ordinalen Messungen die unbeobachteten metrischen Werte zur Verfügung stünden. Da sich solche Abweichungen kaum analytisch bestimmen lassen, erfolgt die Untersuchung meist mit Monte-Carlo-Simulationen. Dazu werden mit dem Computer Zufallszahlen erzeugt, deren Wahrscheinlichkeitsverteilungen den Vorgaben eines linearen Modells entsprechen. In einem zweiten Schritt werden die Zufallszahlen mit Hilfe eines Schwellenwertmodells in ordinale Kategorien transformiert. Da nun sowohl die ursprünglichen metrischen Zufallszahlen vorliegen als auch ordinale Messungen dieser Werte, ist ein direkter Vergleich von Parameterschätzungen auf der Basis der ordinalen wie der metrischen Werte möglich.

Die Untersuchung der Auswirkungen von Mißachtungen des Meßniveaus ist nicht die einzige Anwendungsmöglichkeit von Schwellenwertmodellen. Zunehmend werden sie auch als Ergänzung herkömmlicher linearer Modelle eingesetzt. Mit ihrer Hilfe wird dann versucht, die Momente der unbeobachteten metrischen Größe aus den ordinalen Daten zu schätzen. Auf solche Erweiterungen linearer Modelle beziehen sich die Aussagen, nach denen inzwischen ordinales Skalenniveau bei der Anwendung linearer Modelle explizit berücksichtigt werden kann.

Nun wäre es ausgesprochen erstaunlich, wenn das in ordinalen Daten enthaltene Ausmaß an Informationen einfach in das metrischer Informationen vermehrt werden könnte. Tatsächlich gelingt dies nicht ohne zusätzliche Annahmen. Bei der Anwendung von Schwellenwertmodellen zur Schätzung der Momente der unbeobachteten metrischen Variablen ist es so notwendig, Annahmen über die Verteilung dieser Variablen zu machen. Üblicherweise wird hier unterstellt, daß die Variablen normalverteilt sind.[6] In Abbildung 2 ist die Konsequenz dieser Annahme graphisch verdeutlicht. Die Schwellenwerte τ_1 bis τ_3 zerteilen die Fläche unter der Normalverteilungskurve $f(Y^*)$ in vier Bereiche. Die Flächen der Teilbereiche entsprechen dann den Wahrscheinlichkeiten, mit der Realisationen der metrischen Variable Y^* auf der Beobachtungsebene in die Kategorie A, B, C oder D fallen.

6 Theoretisch lassen sich auch andere Verteilungsformen annehmen. *Karl Jöreskog* meinte hierzu während des Frühjahrsseminars 1993, daß Monte-Carlo-Studien gezeigt hätten, daß die Schätzungen auch bei Verletzung der Normalverteilungsannahme recht robust seien.

Abbildung 2: Normalverteilungsannahme der unbeobachteten metrischen Variable



Die Parameter eines linearen Modells können dann in einem dreistufigen Schätzverfahren bestimmt werden.[7] In der ersten Stufe werden für jede ordinale Variable die Positionen der Schwellenwerte über die relativen Häufigkeiten der Besetzungen in den Kategorien berechnet. Auf der Basis dieser Schätzungen werden in der zweiten Stufe die Momente der unbeobachteten metrischen Variablen sowie deren Varianzen und Kovarianzen geschätzt. In der dritten Stufe werden schließlich aus den Momenten die Modellparameter des linearen Modells ermittelt. Die Varianzen und Kovarianzen der auf der zweiten Stufe geschätzten Momente werden dabei zur Optimierung der Schätzung in der dritten Stufe und zur Bestimmung von Standardschätzfehlern und des Modellfits verwendet.

Anhand von Abbildung 2 läßt sich verdeutlichen, daß die Mittelwerte und Varianzen der unbeobachteten metrischen Variablen Y^* in der Regel nicht unabhängig von den Schwellenwerten bestimmt werden können. Mittelwertsveränderungen bedeuten eine Verschiebung der Normalverteilungskurve entlang der horizontalen Achse. Solche Veränderungen können durch entsprechendes Verschieben der Schwellenwerte kompensiert werden. Veränderungen der Varianzen würden sich in der Abbildung als Stauchung oder Streckung der horizontalen Achse auswirken. Auch hier gilt wieder, daß sich

7 Ein solches dreistufiges Vorgehen, wie es z.B. in PRELIS und LISREL realisiert ist, erleichtert die Handhabbarkeit der Schätzung. Denkbar ist auch die simultane Schätzung aller Größen.

solche Veränderungen durch entsprechendes Anpassen der Positionen der Schwellenwerte ausgleichen lassen.

In LISREL bzw. PRELIS und in EQS werden die Mittelwerte der unbeobachteten metrischen Variablen daher aus Identifikationsgründen auf null fixiert.[8] Die Varianzen der metrischen Variablen werden auf den Wert Eins fixiert. Anstelle der Kovarianzen werden daher aus den bivariaten Häufigkeitstabellen der ordinalen Variablen (polychorische) Korrelationen der unbeobachteten metrischen Variablen geschätzt. Falls die Schwellenwerte a priori bekannt sind, lassen sich auch die Mittelwerte und Varianzen der unbeobachteten metrischen Variablen identifizieren. Es ist außerdem möglich, Mittelwerts- und Varianzunterschiede zwischen den Variablen zu identifizieren, wenn (z.B. bei Paneldaten) Schwellenwerte für verschiedene Variablen gleichgesetzt werden können.[9]

Bei diesen Anwendungen von Schwellenwertmodellen wird davon ausgegangen, daß die Schwellenwerte Parameter sind, die für alle Fälle einer Stichprobe gleich sind. Bei sozialwissenschaftlichen Umfragedaten ist es jedoch auch denkbar, daß bei Einstellungsmessungen sprachlich gekennzeichnete ordinale Antwortkategorien nicht bei allen Befragten eine identische Bedeutung haben. So mag es Personen geben, die in ihren Äußerungen negative Urteile vermeiden. Andere mögen grundsätzlich zu extremen Urteilen neigen. Stark positive bzw. negative Antworten können von diesen Personen auch dann gewählt werden, wenn der 'wahre' Wert eher in der Mitte liegt. Übersetzt in die Logik von Schwellenwertmodellen bedeuten solche unterschiedlichen Antworttendenzen, daß sich die Schwellenwerte von Person zu Person unterscheiden. Bei denjenigen, die negative Urteile vermeiden, sind die Schwellenwerte in Richtung des negativen Pols verschoben; Tendenzen zu extremen Urteilen lassen sich als Zusammenrücken der Schwellenwerte in die Nähe des Mittelpunkts der metrischen Skala modellieren. Ob solche individuellen Variationen die Möglichkeit der Anwendung linearer Modelle auf ordinale Daten einschränken, ist m.W. bisher nicht untersucht worden. Ich möchte daher im folgenden über eine kleine Monte-Carlo-Studie berichten, die hierzu erste Antworten gibt.

---

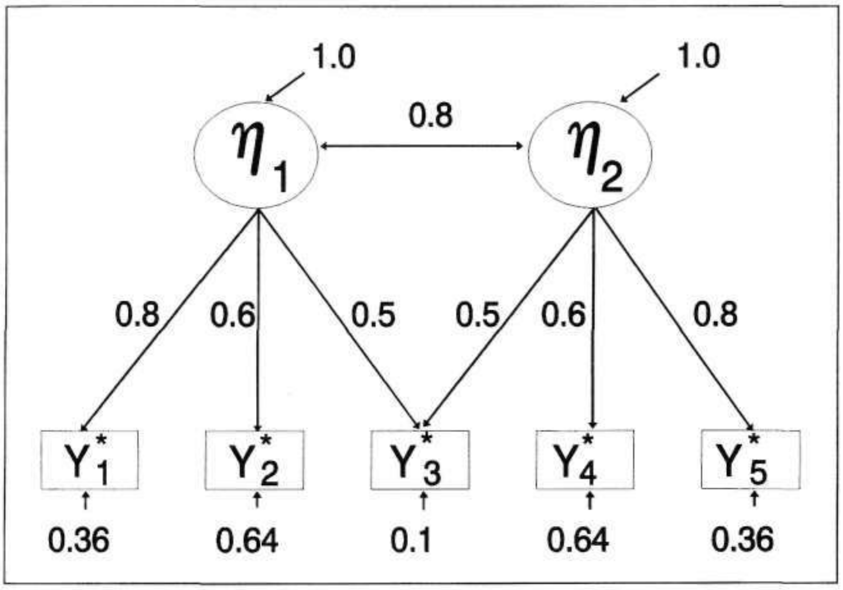
8 *Schepers und Arminger* (1992, S. 9) fixieren dagegen in ihrem Programm MECOSA den Wert des ersten Schwellenwerts auf null.

9 Wenn ein Modell neben den abhängigen ordinalen Variablen beobachtbare exogene Variablen enthält, können zudem mit PRELIS 2 die Mittelwerte und Varianzen der unbeobachteten, metrischen abhängigen Variablen auch über die Regressiongewichte von Probit-Regressionen geschätzt werden.

Eine Monte-Carlo-Simulation zu den Auswirkungen individuell variierender Schwellenwerte

Ziel des Monte-Carlo-Experiments ist es, Hinweise darauf zu gewinnen, ob beim Vorliegen ordinaler Daten (im Sinne der ungenauen Messung metrischer Variablen) lineare Modelle erfolgreich angewendet werden können und ob es dabei notwendig ist, die Ordinalität der Daten durch vorgelagerte Schwellenwertmodelle explizit zu berücksichtigen. Es soll nicht von vornherein angenommen werden, daß die Schwellenwerte für alle Fälle einer Stichprobe gleich sind.

Abbildung 3: Pfaddiagramm des datengenerierenden Modells (Ml)



Ausgangspunkt des Experiments ist das in Abbildung 3 wiedergegebene Pfaddiagramm eines 2-Faktorenmodells mit fünf Indikatoren. Aus den Werten für die Ladungen der Indikatoren auf die Faktoren und den Varianzen und Kovarianzen der Faktoren und der Meßfehler läßt sich die in Tabelle 2 wiedergegebene Varianz/Kovarianz-Matrix der fünf Indikatoren ableiten. Mit Hilfe eines Zufallszahlengenerators wurden 1000 normalverteilte Realisationen der fünf Variablen erzeugt, die diese modellimplizierte Varianz/Kovarianz-Matrix genau erfüllen.[10]

10 Die Generierung der Populations- und Stichprobendaten erfolgte mit SPSS 4 für das Betriebssystem OS/2.

Tabelle 2: Varianzen und Kovarianzen der fünf Indikatoren des Faktorenmodells aus Abbildung 3

Variable	Y_1^*	Y_2^*	Y_3^*	Y_4^*	Y_5^*
Y_1^*	1.000				
Y_2^*	0.480	1.000			
Y_3^*	0.720	0.540	1.000		
Y_4^*	0.384	0.288	0.540	1.000	
Y_5^*	0.512	0.384	0.720	0.480	1.000

Aus dieser Population wurden über einfache Zufallsauswahlen (mit Zurücklegen) 20 Stichproben von jeweils 1000 Fällen gezogen. Die metrischen Stichprobendaten der fünf Variablen sind anschließend in der ersten *Kategorisierung mit festen Schwellenwerten* in ordinale Variablen mit jeweils vier Ausprägungen transformiert worden. Während für die ersten drei Variablen Y_1^* bis Y_3^* die Schwellenwerte so gewählt wurden, daß die ordinalen Ergebnisvariablen Y_1 bis Y_3 unimodal symmetrisch verteilt sind, ist die vierte Variable nach der Kategorisierung schief und die letzte u-förmig verteilt. Tabelle 3 gibt für diese Kategorisierung der metrischen Variablen die Schwellenwerte und die daraus folgenden Wahrscheinlichkeitsverteilungen der ordinalen Variablen wieder.

Tabelle 3: Schwellenwerte und resultierende Wahrscheinlichkeitsverteilung der ordinalen Variablen

Variable:	Y_1	Y_2	Y_3	Y_4	Y_5
τ_1	-1.0	-1.0	-1.0	0.0	-0.5
τ_2	0.0	0.0	0.0	1.0	0.0
τ_3	1.0	1.0	1.0	1.5	0.5
$\pi(Y_i=1)$	15.9 %	15.9 %	15.9 %	50.0 %	30.9 %
$\pi(Y_i=2)$	34.1 %	34.1 %	34.1 %	34.1 %	19.1 %
$\pi(Y_i=4)$	34.1 %	34.1 %	34.1 %	9.2 %	19.1 %
$\pi(Y_i=5)$	15.9 %	15.9 %	15.9 %	6.7 %	30.9 %

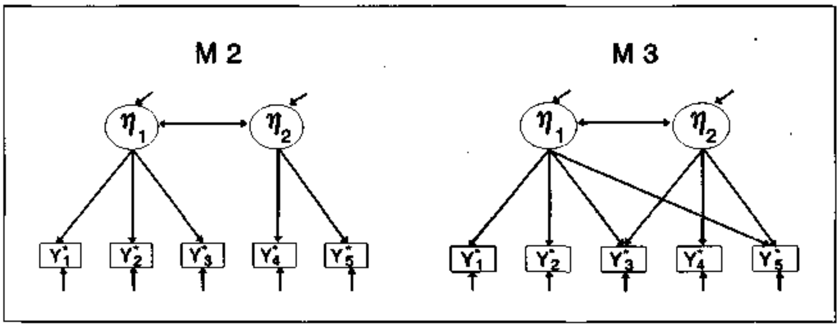
Zusätzlich erfolgte eine zweite *Kategorisierung mit variierenden Schwellenwerten*, bei der die Position der Schwellenwerte in einem Bereich von ± 0.2 zufällig um die festen Schwellenwerte der ersten Kategorisierung streuen. Zur Realisierung der variierenden Schwellenwerte wurden gleichverteilte Zufallszahlen im Wertebereich von -0.2 bis

+0.2 erzeugt. Bei der Kategorisierung der metrischen Variablen eines Falles wurden dann jeweils solche Zufallszahlen auf die Schwellenwerte aus Tabelle 3 aufaddiert.

Für jede der zweimal 20 Stichproben von 5 kategorialen Variablen mit je 1000 Fällen sind mit PRELIS 2 ***(Jöreskog und Sörbom*** 1993a) die *Varianzen und Kovarianzen der fünf kategorialen Variablen* berechnet worden. Die Ordinalität der Daten wurde hier also ignoriert. Für die spätere WLS-Schätzung von linearen Strukturgleichungsmodellen wurden außerdem die (asymptotischen) Varianzen und Kovarianzen der Varianzen und Kovarianzen geschätzt. Um zu prüfen, ob die explizite Berücksichtigung des Meßniveaus bessere Ergebnisse bringt, sind mit mit dem Programm zusätzlich in jeder Stichprobe die *polychorischen Korrelationen* und deren (asymptotische) Varianzen und Kovarianzen geschätzt worden.

Die Varianzen und Kovarianzen oder aber die polychorischen Korrelationen werden im letzten Analyseschritt dazu benutzt, die Parameter konfirmatorischer Faktorenanalysen für die fünf Variablen zu schätzen. Eingesetzt wurde hierzu das Programm LISREL 8 ***(Jöreskog und Sörbom*** 1993b). Neben dem *korrekten Modell* (Ml) aus Abbildung 3 wurden zusätzlich jeweils zwei weitere Modelle geschätzt, die eine abweichende Struktur postulieren. Bei *dem fehlspezifizierten Modell* (M2) wurde die Ladung des dritten Indikators auf den zweiten Faktor (λ_{32}) ausgelassen. Beim *überparametrisierten Modell* (M3) wurde dagegen eine zusätzliche (überflüssige) Ladung des fünften Indikators auf den ersten Faktor (λ_{51}) spezifiziert (Abbildung 4).

Abbildung 4: Pfaddiagramme des fehlspezifizierten Modells (M2) und des überparametrisierten Modells (M3)

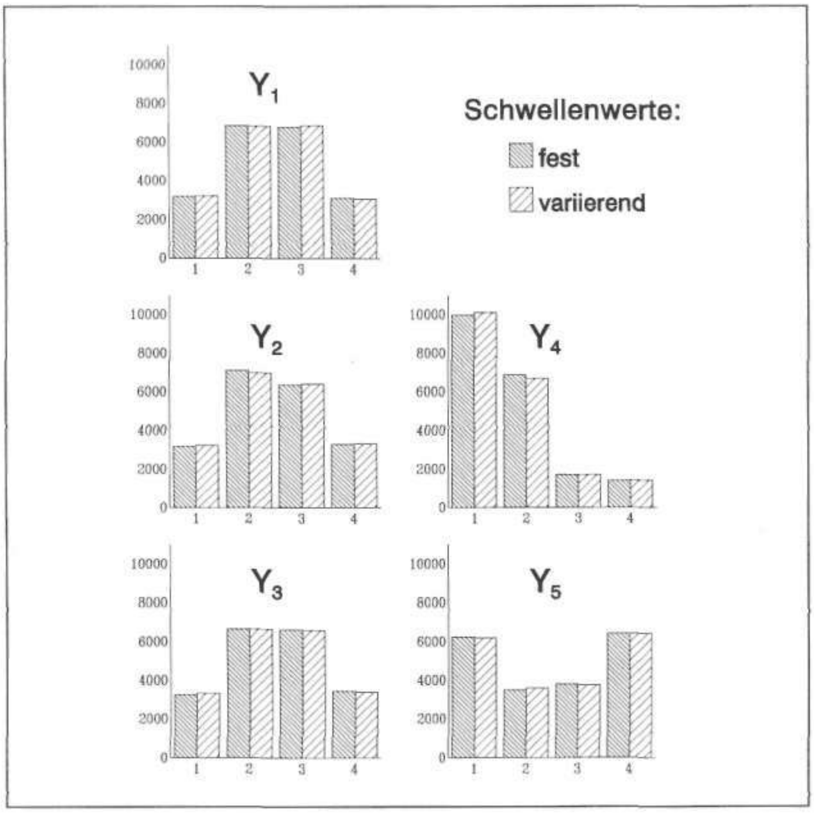


Insgesamt ergibt sich also ein dreifaktorielles Untersuchungsdesign, bei dem zwei Arten von Schwellenwertmodellen, zwei Analysemethoden und drei Analysemodelle unterschieden werden. Alle Ergebnisse in den 12 Zellen basieren auf den gleichen 20 Ursprungsstichproben von je 1000 Fällen und fünf ursprünglich metrischen Variablen.

Ergebnisse

In Abbildung 5 sind die über alle 20 Stichproben kumulierten Häufigkeitsverteilungen der sechs Indikatoren wiedergegeben. Die Verteilungen entsprechen ziemlich genau den in Tabelle 3 wiedergegebenen theoretischen Verteilungen. Auffallend sind die geringen Unterschiede zwischen den beiden Kategorisierungsmethoden. Dies liegt daran, daß die Erwartungswerte bei variierenden Schwellen mit den festen Schwellenwerten zusammenfallen. Tabelle 4 zeigt, daß sich die einzelnen Werte trotz der fast identischen Häufigkeitsverteilungen bei gut 9 % der insgesamt 100 000 Realisationen unterscheiden.

Abbildung 5: Häufigkeitsverteilung der kategorialen Variablen über alle Stichproben



Betrachten wir zunächst die Schätzung der Korrelationen der unbeobachteten metrischen Variablen. In Tabelle 5a sind die Ergebnisse für die Schätzungen der polychorischen

Tabelle 4: Kreuzklassifizierung von festen und variierenden Schwellenwerten über alle Variablen und Stichproben

feste Schwellen	variierende Schwellen 1	2	3	4
1	24547	1563		
2	1338	27680	1800	
3		1829	22256	1283
4			1270	16434

Tabelle 5a: Schätzung der Korrelationen der unbeobachteten metrischen Variablen (polychorische Korrelationen)
a: bei festen Schwellenwerten

Variablenpaar	ρ	Bias	MSE	AV (PRELIS 2)
Y_1, Y_2	0.480	−0.044	.002457	.000938
Y_1, Y_3	0.720	−0.002	.000529	.000358
Y_1, Y_4	0.384	−0.004	.001092	.001276
Y_1, Y_5	0.512	−0.014	.000800	.000890
Y_2, Y_3	0.540	−0.015	.000741	.000770
Y_2, Y_4	0.288	+0.012	.000819	.001247
Y_2, Y_5	0.389	−0.012	.001150	.001105
Y_3, Y_4	0.540	−0.014	.001185	.000856
Y_3, Y_5	0.720	+0.006	.000386	.000369
Y_4, Y_5	0.480	−0.003	.001078	.001007
Mittelwert	-	−0.009	.001024	.000881

Korrelationen bei festen Schwellenwerten wiedergegeben. Aufgeführt sind neben den Zielgrößen (ρ) die über die 20 Stichproben berechneten Mittelwerte der Differenzen zwischen Zielgrößen und Schätzungen (Bias), die durchschnittlichen quadrierten Abweichungen zwischen Zielgrößen und Schätzungen (Mean Squared Error: MSE) und die Durchschnittswerte der von PRELIS 2 berechneten Schätzungen der asymptotischen Varianzen der Korrelationen (AV). Die asymptotischen Varianzen (und Kovarianzen) werden bei der WLS-Schätzung linearer Modelle als Gewichte verwendet. Sowohl

Unter- als auch Überschätzungen können daher zu Fehleinschätzungen der Modellanpassung führen.

Die durchschnittlichen Fehler sind recht gering. Es läßt sich nicht feststellen, daß die Fehler unterschiedlich ausfallen, wenn eine der beteiligten Variablen schief (Y_4) oder u-förmig (Y_5) verteilt ist. Der Spalte mit den durchschnittlichen quadrierten Abweichungen ist zu entnehmen, daß auch die Streuung der Schätzungen nicht sehr hoch ist. Rechnet man aus der Streuung den Anteil der Verzerrung (Bias) heraus, ergibt sich als Schätzung der durchschnittlichen (mit n-1 gewichteten) Varianzen der Korrelationen ein Wert von 0.000993. Dieser Wert ist kaum höher als die in der letzten Spalte wiedergegebene durchschnittliche Schätzung der Varianzen der Korrelationen (0.000881).

Tabelle 5b: Schätzung der Korrelationen der unbeobachteten metrischen Variablen (polychorische Korrelationen)
b: bei variierenden Schwellenwerten

Variablen-paar	ρ	Bias	MSE	AV (PRELIS 2)
Y_1, Y_2	0.480	−0.031	.001573	.000897
Y_1, Y_3	0.720	−0.009	.006296	.000386
Y_1, Y_4	0.384	−0.013	.001177	.001295
Y_1, Y_5	0.512	−0.009	.000736	.000864
Y_2, Y_3	0.540	−0.025	.001208	.000769
Y_2, Y_4	0.288	−0.002	.000965	.001270
Y_2, Y_5	0.389	−0.013	.001075	.001084
Y_3, Y_4	0.540	−0.014	.001232	.000843
Y_3, Y_5	0.720	+0.002	.000377	.000363
Y_4, Y_5	0.480	−0.014	.001667	.001018
Mittelwert	-	−0.012	.001631	.000879

In Tabelle 5b sind die entsprechenden Werte für die Schätzungen bei zufällig variierenden Schwellenwerten wiedergegeben. Auch hier gilt, daß die durchschnittlichen Fehler sehr gering sind. Verglichen mit den Schätzungen bei festen Schwellenwerten ist die Verzerrung (Bias) aber geringfügig höher. Erkennbar höher ist die durchschnittliche quadrierte Abweichung (MSE). Wird auch hier wieder die Verzerrung herausgerechnet, ergibt sich als Schätzung für die durchschnittliche Varianz ein Wert von 0.001565. Der

Durchschnittswert der Schätzungen dieser Größe durch PRELIS 2 (in der Spalte 'AV') fällt mit einem Wert von 0.000879 geringer aus.

Bei der Interpretation dieser Ergebnisse ist Vorsicht angeraten. So bleibt es unklar, ob die Unterschiede in der Höhe der Verzerrung zufällig oder systematisch sind. Die relativ gesehen größeren Abweichungen bei den Streuungen legen schon eher die Vermutung nahe, daß über die Untersuchungseinheiten variierende Schwellenwerte die effektiven Standardschätzfehler vergrößern.

Die polychorischen Korrelationen bilden die Datenbasis für die Schätzung der Modellparameter, wenn die Ordinalität der Daten explizit berücksichtigt werden soll. Wird die Ordinalität der Daten dagegen ignoriert, erfolgt die Modellschätzung auf der Basis von Varianzen und Kovarianzen der kategorialen Variablen, wobei den Kategorien Zahlen von 1 bis 4 zugeordnet sind.

In Tabelle 6 sind die Ergebnisse der Modellanpassungen für alle 12 Zellen des Monte-Carlo-Simulationsexperiments zusammengefaßt. Aufgeführt ist jeweils der über die 20 Stichproben gemittelte Wert des Chiquadrat-Differenzentests ($\bar{\chi}^2$) sowie die geschätzte Varianz ($\hat{\sigma}^2$) dieser Werte. Bei einem korrekt spezifizierten Modell und angemessenem Schätzverfahren sollte die Testgröße asymptotisch chiquadratverteilt sein. Da der Erwartungswert einer Chiquadratverteilung mit der jeweiligen Anzahl der Freiheitsgrade übereinstimmt und die Varianz das Zweifache der Zahl der Freiheitsgrade ist, ist bei der Schätzung des den Daten zugrundeliegenden Modells aus Abbildung 3 mit einem durchschnittlichen Chiquadratwert von 3 und einer Varianz von 6 zu rechnen. Für das überparametrisierte Modell (M3) sollte entsprechend der durchschnittliche Chiquadratwert 2 und die Varianz 4 betragen. [11] Beim fehlspezifizierten Modell (M2) sollten schließlich die durchschnittlichen Chiquadratwerte deutlich höher als 4 sein und somit eine schlechte Modellanpassung signalisieren. Außerdem ist - verglichen mit korrekt spezifizierten Modellen - mit einer größeren Streuung der Testgröße zu rechnen.

Die Werte in Tabelle 6 zeigen sehr deutlich, daß das fehlspezifizierte Modell (M2) tatsächlich im Durchschnitt deutlich höhere Werte aufweist als die beiden anderen Modelle (Ml und M3). Dies gilt unabhängig davon, ob die Schwellenwerte fest sind oder zufällig variieren, und auch unabhängig davon, ob die Analyse auf der Basis polychorischer Korrelationen erfolgt oder die Ordinalität der Messungen ignoriert wird. Es ist also möglich, ein fehlspezifiziertes Modell von einem korrekt spezifizierten Modell zu unterscheiden. Während fehlspezifizierte Modelle aufgrund ihrer hohen Chiquadrat-

11 Der zusätzlich geschätzte Modellparameter λ_{51} sollte dabei nicht signifikant von null verschieden sein.

Tabelle 6: Güte der Modellschätzungen (Chiquadrat-Differenzen-Test)

Schwellenwerte	Methode	Modell	df	$\bar{\chi}^2$	$\hat{\sigma}^2(\chi^2)$
fest	Analyse polychorischer Korrelationen	korrekt spezifiziert	3	4.17	10.05
		überparametrisiert	2	2.16	5.46
		fehlspezifiziert	4	11.26	32.08
	Analyse von Varianzen und Kovarianzen	korrekt spezifiziert	3	5.01	12.12
		überparametrisiert	2	3.00	5.74
		fehlspezifiziert.	4	12.47	35.94
variierend	Analyse polychorischer Korrelationen	korrekt spezifiziert	3	3.40	3.77
		überparametrisiert	2	2.34	2.86
		fehlspezifiziert	4	14.30	30.96
	Analyse von Varianzen und Kovarianzen	korrekt spezifiziert	3	4.19	6.48
		überparametrisiert	2	2.74	3.95
		fehlspezifiziert	4	15.58	34.09

werte Zweifel an der Angemessenheit der Modellspezifikation wecken, weisen die Anpassungstests für das korrekt spezifizierte Modell und für das überparametrisierte Modell dagegen darauf hin, daß das jeweilige Modell mit den Daten verträglich erscheint.

Wenn das Modell korrekt spezifiziert ist, sind die Werte der Anpassungstests bei variierenden Schwellenwerten eher etwas geringer als bei festen Schwellenwerten. Bei fehlspezifizierten Modellen sind die Chiquadratwerte dagegen eher etwas höher. Verglichen mit den Modellschätzungen auf der Basis polychorischer Korrelationen sind die durchschnittlichen Chiquadratwerte bei der Ignorierung des ordinalen Meßniveaus geringfügig höher. Dies gilt auch für die Streuung der Werte. Allerdings sind die Unterschiede nicht so groß, daß sie Auswirkungen auf die Beurteilungen der Modellanpassung haben. Insgesamt gesehen weisen die Ergebnisse der Modellschätzung darauf hin, daß sich lineare Modelle auch dann erfolgreich anwenden lassen, wenn die Daten ordinal im Sinne ungenauer Erfassungen metrischer Variablen sind. Für die Beurteilung der Modellanpassung scheint es dabei auch nicht unbedingt notwendig zu sein, die Ordinalität der Daten explizit zu berücksichtigen. Inwieweit dies auch für die Parameterschätzung selber gilt, ist aus Tabelle 7a bzw. 7b ersichtlich.

Tabelle 7a: Genauigkeit der Parameterschätzungen (Modell Ml)
a: feste Schwellenwerte

	Methode					
	Analyse polychorischer Korrelationen			Analyse von Varianzen und Kovarianzen		
Parameter	Bias	$\hat{\sigma}$	MSE	Bias	$\hat{\sigma}$	MSE
$\rho_{21}=0.8$	.013	.032	.00116	.014	.036	.00141
$\lambda_{11}=0.8$	-.028	.033	.00182	-.073	.033	.00636
$\lambda_{21}=0.6$	-.032	.022	.00147	-.068	.022	.00508
$\lambda_{31}=0.5$	.096	.089	.01671	.049	.086	.00949
$\lambda_{32}=0.5$	-.090	.090	.01577	-.087	.086	.01464
$\lambda_{42}=0.6$	-.008	.030	.00094	-.092	.025	.00906
$\lambda_{52}=0.8$	.014	.035	.00139	-.055	.036	.00428
Mittelwert	-.005	.047	.00561	-.045	.042	.00719

In der ersten Tabelle (7a) sind die Ergebnisse der Parameterschätzung bei festen Schwellenwerten wiedergegeben. Der Vergleich der Verzerrungen (Bias) der beiden Methoden zeigt, daß die Nichtberücksichtigung der Ordinalität im Durchschnitt deutlich größere Verzerrungen ergibt. Bei der Analyse polychorischer Korrelationen sind die Verzerrungen im Durchschnitt recht gering. Auffallend ist aber, daß die Werte für die einzelnen Modellparameter sehr unterschiedlich ausfallen. Die Ladungen des dritten Indikators (λ_{31} und λ_{32}) werden bei der Analyse polychorischer Korrelationen nicht besser geschätzt als bei der Mißachtung des Meßniveaus. Zurückzuführen ist dies allerdings nicht auf die Schätzmethode, sondern auf die Modellstruktur. Aufgrund der hohen Korrelation der beiden Faktoren ($\rho_{21}=0.8$) gibt es hier ein Multikollinearitätsproblem. Sichtbar wird es auch durch die deutlich höheren Streuungen bei den Schätzungen der beiden Parameter. Die über die 20 Stichproben berechneten Standardschätzfehler ($\hat{\sigma}$) sind hier zwei- bis dreimal so hoch wie bei den übrigen Modellparametern.

Um die Ergebnisse zwischen den beiden Methoden vergleichen zu können, sind in **den** Tabellen 7a und 7b bei Analysen auf der Basis von Varianzen und Kovarianzen die Schätzwerte für die (voll-) standardisierten Ladungen aufgeführt und nicht die ursprünglichen Schätzungen des unstandardisierten Modells.[12] Bei den Faktorladungen ergeben sich deutliche Unterschätzungen der korrekten Werte. (Die hiervon abweichende positive Verzerrung bei λ_{31} ist eine Folge der erwähnten Multikolinearität. Dieses Ergebnis kann vor allem darauf zurückgeführt werden, daß die Korrelationen bei der Mißachtung des Skalenniveaus im Durchschnitt geringer ausfallen als die entsprechenden polychorischen Korrelationen. Da geringere (standardisierte) Ladungen mit höheren Meßfehlervarianzen korrespondieren, werden die Meßfehler also eher überschätzt. Als Folge werden auch Beziehungen zwischen den latenten Variablen eher überschätzt. Diese Überschätzung ist allerdings in diesem Fall praktisch kaum höher als die entsprechende Verzerrung bei der Analyse polychorischer Korrelationen.

Tabelle 7b: Genauigkeit der Parameterschätzungen (Modell Ml)
b: variierende Schwellenwerte

	Methode					
	Analyse polychorischer Korrelationen			Analyse von Varianzen und Kovarianzen		
Parameter	Bias	$\hat{\sigma}$	MSE	Bias	$\hat{\sigma}$	MSE
$\rho_{21}=0.8$	.008	.033	.00111	.011	.036	.00136
$\lambda_{11}=0.8$	-.016	.038	.00166	-.058	.035	.00452
$\lambda_{21}=0.6$	-.027	.023	.00125	-.062	.023	.00437
$\lambda_{31}=0.5$	.004	.117	.01300	-.044	.114	.01431
$\lambda_{32}=0.5$	-.005	.118	.01329	-.003	.115	.01248
$\lambda_{42}=0.6$	-.015	.037	.00153	-.098	.030	.01046
$\lambda_{52}=0.8$	.001	.033	.00106	-.067	.033	.00551
Mittelwert	-.007	.057	.00470	-.046	.055	.00757

Kommen wir nun zu den Ergebnissen bei zufällig variierenden Schwellenwerten. Auch hier zeigt der Vergleich der beiden Methoden die erwähnten Unterschiede: Basiert die Analyse auf polychorischen Korrelationen, sind die durchschnittlichen Verzerrungen **der** geschätzten Modellparameter deutlich geringer. Bei einem genaueren Vergleich der

12 Werden die Parameter der Modelle M1 bis M3 auf der Basis der polychorischen Korrelationen geschätzt, sind standardisierte und unstandardisierte Schätzungen identisch.

beiden Tabellen fällt auf, daß das Multikolinearitätsproblem nicht so stark durchschlägt. Die Ladungen λ_{31} und λ_{32} sind nicht stärker verzerrt als die übrigen Ladungen. Deutlich höher ist hier nur die Streuung der Schätzwerte. Generell sind die Streuungen der Parameterschätzungen jedoch kaum höher als bei den Stichproben mit festen Schwellenwerten. Dieses Ergebnis legt eine Interpretation nahe, nach der die zufällig variierenden Schwellenwerte faktisch wie eine Erhöhung der tatsächlichen Meßfehlervarianz wirken. Aufgrund der dadurch etwas geringeren Korrelationen der Indikatoren wird das Multikolinearitätsproblem entschärft.

Die bisher berichteten Ergebnisse weisen darauf hin, daß es bei der Mißachtung der Ordinalität der Daten zwar möglich ist, fehlspezifizierte von korrekten Modellen zu unterscheiden, daß die Schätzung der Modellparameter aber genauer wird, wenn das Meßniveau durch vorgeschaltete Schwellenwertmodelle berücksichtigt wird. Solange man nur an der Modellstruktur interessiert ist, kann hiernach die Analyse auch auf der Basis der Varianzen und Kovarianzen kategorialer Variablen erfolgen.

Tabelle 8: Detailergebnisse des fehlspezifizierten Modells

	feste Schwellenwerte		variierende Schwellenwerte	
	Analyse polychorischer Korrelationen	Analyse von Varianzen	Analyse polychorischer Korrelationen	Analyse von Varianzen
LM(λ_{32})	11.261	12.471	14.297	15.578
$\hat{\sigma}$(LM)	5.664	5.995	5.565	5.838
LM > 4	19 / 20	19 / 20	19 / 20	20 / 20
LM > 7	16 / 20	17 / 20	19 / 20	20 / 20
Maximum	10 / 20	10 / 20	14 / 20	13 / 20
Negative Varianzen	6 / 20	0 / 20	3 / 20	0 / 20

Wenn - was in der Praxis der Regelfall sein dürfte - das ursprünglich spezifizierte Modell nicht auf die Daten paßt, stellt sich aber auch die Frage, ob sich trotz der Mißachtung des Skalenniveaus auch Hinweise darauf finden lassen, in welche Richtung eine etwaige Modellmodifikation gehen sollte. Ein Kriterium, das bei Modellmodifikationen oft verwendet wird, ist der Lagrange-Multiplier- oder Score-Test (LM-Test). [13] 13

13 In der LISREL-Programmausgabe werden univariate LM-Tests als 'Modifikationsindizes' bezeichnet.

Im fehlspezifizierten Modell (M2 aus Abbildung 4) sollte der LM-Test bei der ausgelassenen Ladung λ_{32} durch hohe Werte auf diesen Fehler der Modellspezifikation hinweisen. Bei ansonsten korrekten Modellannahmen ist die Teststatistik asymptotisch chiquadratverteilt. Als Faustregel hat sich eingebürgert, Modifikationsindizes als bedeutsam zu betrachten, die größer als 7 sind. In Tabelle 8 ist neben dem Durchschnittswert und der Standardabweichung der LM-Tests über die 20 Stichproben auch angegeben, in wievielen der Stichproben die Teststatistik größer 4 bzw. größer 7 ist. Der kleinere Wert entspricht einem Signifikanzniveau von etwa 5 %.

Die Ergebnisse sind recht eindeutig. Unabhängig davon, ob die Schwellenwerte fest oder variierend sind, und unabhängig davon, ob polychorische Korrelationen analysiert werden oder die Varianzen und Kovarianzen der kategorialen Variablen, der Durchschnittswert der LM-Test liegt stets deutlich über 10. Bei variierenden Schwellenwerten ist der Durchschnittswert eher noch höher als bei festen Schwellenwerten. Geringfügig höher ist er auch bei der Ignorierung des ordinalen Meßniveaus. In der Tabelle ist weiter aufgeführt, wie oft die LM-Teststatistik für λ_{32}, verglichen mit den Werten für andere Modellparameter maximal ist. Dies trifft in mindestens der Hälfte aller Stichproben zu. Bei der Analyse polychorischer Korrelationen weisen zudem negative Varianzen bei den Meßfehlern in sechs bzw. drei Stichproben darauf hin, daß das Modell nicht korrekt spezifiziert ist. Die Werte in Tabelle 8 lassen den Schluß zu, daß es nicht nur möglich ist, fehlspezifizierte Modelle von korrekten Modellen zu unterscheiden, sondern auch, Hinweise auf die Ursache einer schlechten Anpassung an die Daten zu finden. Anhand der Werte in Tabelle 9 läßt sich prüfen, ob auch umgekehrt gilt, daß der überflüssige Parameter λ_{51} im überparametrisierten Modell (M3 aus Abbildung 4) entdeckt wird.

Aufgeführt ist zum einen der durchschnittliche Schätzwert für diesen Parameter ($\bar{\lambda}_{51}$). Obwohl der Wert in der Grundgesamtheit null ist, ergeben sich bei allen vier Versuchsbedingungen im Mittel negative Werte. Zu beachten ist allerdings, daß die von LISREL 8 berechneten Standardschätzfehler im Durchschnitt sehr hoch sind. Infolgedessen sind die Absolutbeträge der T-Werte, die sich durch Division der Parameter durch ihre Standardschätzfehler ergeben, im Durchschnitt eher klein. Da die T-Werte asymptotisch normalverteilt sein sollen, kann mit ihnen die Signifikanz eines Parameters geschätzt werden. Tatsächlich tritt in keiner einzigen Stichprobe ein signifikanter Wert auf. Die Überparametrisierung der Modelle wird danach also genauso entdeckt wie das Nichtberücksichtigen von Parametern.

Aus der letzten Zeile von Tabelle 9 wird ersichtlich, daß das überparametrisierte Modell nicht in allen 20 Stichproben geschätzt werden konnte. Vor allem bei der Analyse von Varianzen und Kovarianzen brach der Schätzalgorithmus nach Erreichen der maxi-

Tabelle 9: Detailergebnisse des überparametrisierten Modells

	feste Schwellenwerte		variierende Schwellenwerte	
	Analyse polychorischer Korrelationen	Analyse von Varianzen	Analyse polychorischer Korrelationen	Analyse von Varianzen
$\bar{\lambda}_{s1}$	−0.331	−0.368	−0.290	-0.171
durchschnittlicher Standardfehler	0.318	0.358	0.688	0.331
$\|\bar{T}\|$	0.596	0.647	0.739	0.797
Streuung der T-Werte	0.308	0.361	0.488	0.439
Signifikanz (α=5%)	0 / 17	0 / 13	0 / 19	0 / 17
Keine Konvergenz	3 / 20	7 / 20	1 / 20	3 / 20

malen Iterationszahl ab, ohne zu einer Lösung konvergiert zu haben. Das Nichterreichen der Konvergenz kann ebenfalls als ein Hinweis interpretiert werden, daß das spezifizierte Modell möglicherweise nicht optimal ist. Allerdings ist es in solchen Fällen oft recht schwer, die Ursache des Problems zu lokalisieren.

Diskussion

Die in diesem Beitrag vorgestellte Monte-Carlo-Simulationsstudie hat gezeigt, daß es . möglich ist, lineare Strukturgleichungsmodelle bei ordinalen Daten anzuwenden. Selbst wenn man die Ordinalität der Daten ignoriert und die Varianzen und Kovarianzen der durchnumerierten kategorialen Variablen analysiert, können korrekt spezifizierte lineare Modelle identifiziert und Fehlspezifikationen als auch Überparametrisierungen entdeckt und deren Ursachen lokalisiert werden. Wird die Ordinalität der Daten mit Hilfe von vorgeschalteten Schwellenwertmodellen im Analysemodell berücksichtigt, können zudem die Parameterwerte eines korrekt spezifizierten Modells recht genau geschätzt werden. Bislang noch unbekannt war, daß dies selbst dann gilt, wenn die Schwellenwerte nicht - wie üblicherweise angenommen - konstant sind, sondern zufällig über die Untersuchungseinheiten variieren.

Dieses an sich sehr positive Ergebnis muß jedoch relativiert werden. Da es auf einer Monte-Carlo-Simulation beruht, bezieht es sich auf die im Untersuchungsdesign festgelegte spezifische Datenkonstellation. Verallgemeinerungen können sich als falsch erweisen. Es scheint daher ratsam, sich die Besonderheiten dieser Untersuchung vor Augen zu führen.

Diese betreffen bereits das lineare Modell, mit dem die Daten erzeugt worden sind. Es handelt sich hier um ein relativ einfaches Faktorenmodell mit nur fünf Indikatoren. Offen bleibt, ob die Ergebnisse auch so eindeutig ausgefallen wären, wenn eine komplexere Struktur untersucht worden wäre, bei der Meßfehler korrelieren, die meisten Indikatoren durch mehrere Faktoren erklärt würden und die latenten Variablen durch ein differenziertes Kausalmodell verbunden wären. Da das datengenerierende Modell nur wenige Variablen umfaßt, ist auch das Ausmaß der untersuchten Fehlspezifikation nicht sehr hoch.

Die zunächst metrischen Indikatoren sind in ordinale Variablen mit jeweils vier Kategorien transformiert worden. Es ist eher unwahrscheinlich, daß bei einer höheren Anzahl von Kategorien die Ergebnisse ungünstiger ausfallen. Umgekehrt ist aber durchaus denkbar, daß bei dichotomen oder trichotomen Variablen die Ignorierung des Meßniveaus problematischer ist. Insofern kann die oben erwähnte Aussage von ***Bentler und Chou***, nach der bei weniger als vier Kategorien die Ordinalität von Daten nicht ignoriert werden sollte, aufgrund meiner Ergebnisse nicht abgeschwächt werden.

In dieser kleinen Monte-Carlo-Studie wurde auch nicht untersucht, ob sich die Ergebnisse ändern, wenn die metrischen Variablen nicht normalverteilt sind. Die Annahmen des klassischen Schwellenwertmodells wurden nur dahingehend abgeschwächt, daß die Schwellenwerte in einer Versuchsbedingung zufällig um jeweils einen Wert schwankten. Diese Einführung variierender Schwellenwerte kann nur ein erster Schritt einer tiefergehenden Analyse sein. Insbesondere ist noch zu klären, welche Auswirkungen es hat, wenn die Schwellenwerte nicht zufällig, sondern systematisch variieren und diese Systematik mit den Variablen des Modells korreliert.

Trotz dieser Einschränkungen sollte doch deutlich geworden sein, daß ein prinzipielles Ablehnen der Anwendung linearer Modelle bei ordinalen Daten durch den pauschalen Verweis auf das fehlende metrische Skalenniveau nicht angebracht ist. Genauso verkehrt wäre es aber, daraus umgekehrt zu schließen, daß alle ordinalen Daten mit linearen Modellen analysiert werden können. Die zentrale Voraussetzung der Anwendung linearer Modelle ist, daß die ordinalen Daten als ungenaue Erfassung metrischer Größen aufgefaßt werden können. Diese Voraussetzung kann meiner Ansicht nach weder als grundsätzlich nicht erfüllbar noch als grundsätzlich erfüllt gelten, sondern muß in jedem Einzelfall neu durchdacht werden.

Literatur

Bentler, Peter M.; Chou Chih-Ping, 1987:
Practical Issues in Structural Modeling. In: Sociological Methods and Research, 16, S. 78-117.

Bentler, Peter M.; Wu, Eric J.C., 1993:
EQS/Windows User's Guide. Los Angeles: BMDP Statistical Software, Inc.

Jöreskog, Karl G.; Sörbom, Dag 1993a:
PRELIS 2: User's Reference Guide. Chicago: Scientific Software International, Inc.

Jöreskog, Karl G.; Sörbom, Dag 1993b:
LISREL 8: 2 User's Reference Guide. Chicago: Scientific Software International, Inc.

Jöreskog, Karl G.;Sörbom, Dag 1993c:
LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. Chicago: Scientific Software International, Inc.

Kriz, Jürgen 1981:
Methodenkritik empirischer Sozialforschung, Stuttgart: Teubner.

Kromrey, Helmut 1986:
Empirische Sozialforschung. 3. Auflage, Opladen: Leske u. Budrich (UTB).

Schepers, Andreas; Arminger, Gerhard 1992:
MECOSA version 2.0: User Guide, Zürich: SLI-AG.

Schnell, Rainer; Hill, Paul; Esser, Elke 1993:
Methoden der empirischen Sozialforschung. 4. Auflage, München: Oldenbourg.